

Amendments to the Claims:

The following listing of claims will replace all prior versions, and listings, of claims in the application:

1 1. (Currently amended) A method for distributing data items from a particular set of data into
2 a plurality of buckets based on distribution keys associated with said data items, the
3 method comprising the steps of:

AB 4 randomly selecting data items from said particular set of data to produce a
5 sampled set of data items;

6 ~~determining a plurality of ranges based on the distribution keys associated with~~
7 ~~the sampled set of data items;~~

8 determining a range for each bucket of said plurality of buckets based on a
9 number of the distribution keys associated with said sampled set of
10 data items that fall within said range;

11 assigning said ~~plurality of range~~[[s]] to each bucket of said plurality of
12 buckets; and

13 distributing each data item in said particular set of data to the bucket that has
14 been assigned the range into which falls the distribution key of the data
15 item.

1 2. (Currently amended) ~~The method of Claim 1 wherein the step of~~ A method for distributing
2 data items from a particular set of data into a plurality of buckets based on distribution
3 keys associated with said data items, the method comprising the steps of:

4 randomly selecting data items ~~from said particular set of data includes~~

5 ~~randomly selecting data items~~ from each subset of a plurality of subsets
6 of said particular set of data[.];

7 determining a plurality of ranges based on the distribution keys associated with
8 the sampled set of data items;
9 assigning said plurality of ranges to said plurality of buckets; and
10 distributing each data item in said particular set of data to the bucket that has
11 been assigned the range into which falls the distribution key of the data
12 item.

AB
1 3. (Original) The method of Claim 2 wherein the step of randomly selecting data items from
2 each subset of a plurality of subsets of said particular set of data includes randomly
3 selecting data items from each partition of a partitioned table.

1 4. (Original) The method of Claim 2 wherein the step of randomly selecting data items from
2 each subset of a plurality of subsets of said particular set of data includes randomly
3 selecting data items from subsets of data, stored in buffers in volatile memory, that
4 represent results of one or more previously performed operations.

1 5. (Original) The method of Claim 1 further comprising the steps of:
2 assigning the plurality of buckets to a plurality of processes; and
3 causing each process of said plurality of processes to perform, in parallel with the
4 other processes of said plurality of processes, an operation on the data items
5 contained in any buckets assigned to the process.

1 6. (Original) The method of Claim 2 further comprising the step of selecting a distinct
2 random seed for each subset of the plurality of subsets of said particular set of data.

1 7. (Currently amended) ~~The method of Claim 1 wherein:~~ A method for distributing data items
2 from a particular set of data into a plurality of buckets based on distribution keys
3 associated with said data items, the method comprising the steps of:
4 randomly selecting data items from said particular set of data to produce a
5 sampled set of data items; wherein
6 the particular set of data is durably stored on a plurality of
7 durable storage units; and
8 the step of randomly selecting data items ~~from said particular~~
9 ~~set of data to produce a sampled set of data items~~
10 includes randomly selecting durable storage units from
11 said plurality of durable storage units and using the data
12 items stored on said randomly selected durable storage
13 units as the sampled set of data items[.];
14 determining a plurality of ranges based on the distribution keys associated with
15 the sampled set of data items;
16 assigning said plurality of ranges to said plurality of buckets; and
17 distributing each data item in said particular set of data to the bucket that has
18 been assigned the range into which falls the distribution key of the data
19 item.

1 8. (Original) The method of Claim 1 wherein the step of randomly selecting data items
2 includes selecting a specified percentage of data items in said particular set of data.

1 9. (Original) The method of Claim 7 wherein the step of randomly selecting data items
2 includes selecting a specified percentage of the plurality of durable storage units that
3 are storing said particular set of data.

1 10. (Original) The method of Claim 8 further comprising the step of receiving, from a user,
2 data that specifies said percentage.

1 11. (Original) The method of Claim 9 further comprising the step of receiving, from a user,
2 data that specifies said percentage.

1 12. (Original) The method of Claim 5 wherein said operation is specified in a database
2 command, the method further comprising receiving with said database command data
3 that indicates how much of said particular set of data to randomly select to produce
4 said sampled set of data items.

1 13. (Currently amended) The method of Claim 1 wherein ~~the step of determining a plurality~~
2 ~~of ranges based on the distribution keys associated with the sampled set of data items~~
3 ~~includes determining~~ said ranges ~~[[that]]~~ contain an approximately equal ~~[[amount]]~~
4 number of distribution keys associated with said sampled set of data items.

1 14. (Currently amended) A computer-readable medium carrying instructions for distributing
2 data items from a particular set of data into a plurality of buckets based on distribution

3 keys associated with said data items, the instructions comprising instructions for
4 performing the steps of:

5 randomly selecting data items from said particular set of data to produce a

6 sampled set of data items;

7 ~~determining a plurality of ranges based on the distribution keys associated with~~

8 ~~the sampled set of data items;~~

9 determining a range for each bucket of said plurality of buckets based on a

10 number of the distribution keys associated with said sampled set of

11 data items that fall within said range;

12 assigning said ~~plurality of range~~[[s]] to each bucket of said plurality of

13 buckets; and

14 distributing each data item in said particular set of data to the bucket that has

15 been assigned the range into which falls the distribution key of the data

16 item.

13 15. (Currently amended) ~~The computer-readable medium of Claim 14 wherein the step of A~~

2 computer-readable medium carrying instructions for distributing data items from a

3 particular set of data into a plurality of buckets based on distribution keys associated

4 with said data items, the instructions comprising instructions for performing the steps

5 of:

6 randomly selecting data items ~~from said particular set of data includes~~

7 ~~randomly selecting data items~~ from each subset of a plurality of subsets

8 of said particular set of data[.];

9 determining a plurality of ranges based on the distribution keys associated with
10 the sampled set of data items;
11 assigning said plurality of ranges to said plurality of buckets; and
12 distributing each data item in said particular set of data to the bucket that has
13 been assigned the range into which falls the distribution key of the data
14 item.

A3
1 16. (Original) The computer-readable medium of Claim 15 wherein the step of randomly
2 selecting data items from each subset of a plurality of subsets of said particular set of
3 data includes randomly selecting data items from each partition of a partitioned table.

1 17. (Original) The computer-readable medium of Claim 15 wherein the step of randomly
2 selecting data items from each subset of a plurality of subsets of said particular set of
3 data includes randomly selecting data items from subsets of data, stored in buffers in
4 volatile memory, that represent results of one or more previously performed
5 operations.

1 18. (Original) The computer-readable medium of Claim 14 further comprising instructions for
2 performing the steps of:
3 assigning the plurality of buckets to a plurality of processes; and
4 causing each process of said plurality of processes to perform, in parallel with the
5 other processes of said plurality of processes, an operation on the data items
6 contained in any buckets assigned to the process.

1 19. (Original) The computer-readable medium of Claim 15 further comprising instructions for
2 performing the step of selecting a distinct random seed for each subset of the plurality
3 of subsets of said particular set of data.

AB 1 20. (Currently amended) ~~The computer readable medium of Claim 14 wherein:~~ A computer-
2 readable medium carrying instructions for distributing data items from a particular set
3 of data into a plurality of buckets based on distribution keys associated with said data
4 items, the instructions comprising instructions for performing the steps of:
5 randomly selecting data items from said particular set of data to produce a
6 sampled set of data items; wherein
7 the particular set of data is durably stored on a plurality of
8 durable storage units; and
9 the step of randomly selecting data items from said particular
10 set of data to produce a sampled set of data items
11 includes randomly selecting durable storage units from
12 said plurality of durable storage units and using the data
13 items stored on said randomly selected durable storage
14 units as the sampled set of data items[.];
15 determining a plurality of ranges based on the distribution keys associated with
16 the sampled set of data items;
17 assigning said plurality of ranges to said plurality of buckets; and

18 distributing each data item in said particular set of data to the bucket that has
19 been assigned the range into which falls the distribution key of the data
20 item.

1 21. (Original) The computer-readable medium of Claim 14 wherein the step of randomly
2 selecting data items includes selecting a specified percentage of data items in said
3 particular set of data.

AB

1 22. (Original) The computer-readable medium of Claim 20 wherein the step of randomly
2 selecting data items includes selecting a specified percentage of the plurality of
3 durable storage units that are storing said particular set of data.

1 23. (Original) The computer-readable medium of Claim 21 further comprising instructions for
2 performing the step of receiving, from a user, data that specifies said percentage.

1 24. (Original) The computer-readable medium of Claim 22 further comprising instructions for
2 performing the step of receiving, from a user, data that specifies said percentage.

1 25. (Original) The computer-readable medium of Claim 18 wherein said operation is
2 specified in a database command, the computer-readable medium further comprising
3 instructions for receiving with said database command data that indicates how much
4 of said particular set of data to randomly select to produce said sampled set of data
5 items.

1 26. (Currently amended) The computer-readable medium of Claim 14 wherein ~~the step of~~
2 ~~determining a plurality of ranges based on the distribution keys associated with the~~
3 ~~sampled set of data items includes determining~~ said ranges ~~[[that]]~~ contain an
4 approximately equal ~~[[amount]]~~ number of distribution keys associated with said
5 sampled set of data items.
